



Infrastructure DIRAC pour les Communautés Pluridisciplinaires

A. Tsaregorodtsev, V. Hamar, M. Sapunov, T. Glatard, S. Camarasu-Pop, R.
Ferreira da Silva, Pascal Girard, P. Gay

► To cite this version:

A. Tsaregorodtsev, V. Hamar, M. Sapunov, T. Glatard, S. Camarasu-Pop, et al.. Infrastructure DIRAC pour les Communautés Pluridisciplinaires. Rencontres Scientifiques France Grilles 2011, Sep 2011, Lyon, France. hal-00653025

HAL Id: hal-00653025

<https://hal.science/hal-00653025>

Submitted on 16 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Infrastructure DIRAC pour les Communautés Pluridisciplinaires

A.Tsaregorodtsev(1), V.Hamar(2), M.Sapunov(3), T.Glatard(4), S.Camarasu-Pop(5), R.Ferreira da Silva(6), P.Girard(7), P.Gay(8)

(1) atsareg@in2p3.fr, Centre de Physique des Particules de Marseille, CNRS

(2) hamar@c ppm.in2p3.fr, Centre de Physique des Particules de Marseille, CNRS et Projet GISELA

(3) sapunov@c ppm.in2p3.fr, Centre de Physique des Particules de Marseille, CNRS

(4) tristan.glatard@creatis.insa-lyon.fr, Université de Lyon, CNRS, INSERM, CREATIS

(5) sorina.camarasu@creatis.insa-lyon.fr, Université de Lyon, CNRS, INSERM, CREATIS

(6) rafael.silva@creatis.insa-lyon.fr, Université de Lyon, CNRS, INSERM, CREATIS

(7) pierre.girard@in2p3.fr, Centre de Calcul de l'IN2P3, Lyon

(8) pierre.gay@u-bordeaux1.fr, Université Bordeaux I

Overview:

An increasing number of scientific applications from various domains necessitate the use of large volumes of computing resources. Such resources are provided by computational grid infrastructures but also by commercial clouds and standalone clusters. However, adaptation of the applications for running on the grid is a complicated task. It requires creating specialized portals for managing the application workflows to hide the complexity of the underlying grid resources. The user training becomes essential for their successful use of the grid infrastructures as well.

The DIRAC project brings solutions to the above problems. It provides a Framework for building distributed computing systems giving access to the grids, commercial clouds and standalone clusters. The DIRAC Workload Management System introduced the concept of pilot jobs to solve the problem of interoperability of heterogeneous distributed computing infrastructures. It also allows managing policies and priorities of large user communities sharing common resources. Redundant services and a special system for recovering failures of any grid operation proved to be essential in unstable grid environments. DIRAC File Catalog, storage access clients and automated data distribution system complement the set of services making DIRAC a complete solution for managing distributed computing resources. Its modular architecture is designed for easy extension as needed by specific applications. In the whole, the DIRAC project provides a complete stack of the grid middleware based on the architecture and design principles different from the conventional middleware.

A service based on the DIRAC project is in operation in the Lyon Computing Center. It is used for training courses for the users of the French NGI infrastructure but it is also used by the users from various application domains for accessing their grid resources. In this contribution the DIRAC service installation destined for the French NGI users is described together with the experience of its use for the user tutorials and for running user application.

Enjeux scientifiques, besoin de la grille :

De nombreuses applications scientifiques provenant de disciplines différentes ont besoin de calculs intensifs. Ces calculs nécessitent l'utilisation de ressources de calcul que l'on trouve aujourd'hui dans les grilles de calcul, dans les nuages commerciaux et dans les grappes autonomes. Certaines communautés scientifiques internationales ont accès à des ressources informatiques faisant appel à différentes technologies et provenant de différents domaines administratifs. Le regroupement de toutes ces ressources dans un système unique pouvant être consulté par les utilisateurs à travers une interface uniforme devient un défi majeur.

Les utilisateurs des grilles de calcul sont rarement experts en informatique. Pour cette raison ils ont besoin d'outils qui facilitent l'accès à la grille et cachent la complexité de l'infrastructure sous-jacente. Les tutoriels et le travail collaboratif des utilisateurs et des experts en calculs distribués constituent des éléments essentiels pour la réussite des nouvelles

applications que l'on déploie sur la grille. Souvent les applications ont besoin de services spécifiques pour gérer les enchaînements complexes des tâches qui les composent. Le développement de nouveaux services doit être supporté par une architecture ouverte dans intergiciel ce qui n'est pas toujours le cas avec l'intergiciel standard de la grille EGI.

Développements, déploiement sur la grille :

Le projet DIRAC fournit des logiciels pour créer des systèmes de calcul distribué conformes aux standards de sécurité de la grille de calcul. DIRAC a été développé pour l'expérience en Physique des Hautes Energies LHCb au LHC, CERN. Il constitue la base du modèle de calcul de LHCb et assure la mise à disposition de la performance nécessaire pour cette grande collaboration de plus de 500 physiciens. Aujourd'hui, le projet a été généralisé pour une utilisation dans tous les domaines scientifiques et plus seulement pour la Physique des Hautes Energies.

En ce qui concerne le problème de l'hétérogénéité des ressources de calcul, l'approche la plus souvent appliquée est basée sur la notion de « l'interopérabilité » de différentes infrastructures distribuées. Cette approche a rencontré peu de succès et en pratique limite la flexibilité de l'utilisation des ressources par les applications. Le projet DIRAC emploie le paradigme de surcouche sous forme d'agents-pilotes mobiles déployés sur les nœuds de calcul et qui présentent ces nœuds de façon uniforme à un ordonnanceur central des tâches des utilisateurs (**Fig.1**). Cette méthode permet non seulement de masquer l'hétérogénéité des ressources de calcul mais elle permet aussi de résoudre d'autres problèmes typiques des grilles de calcul. D'abord, l'efficacité du travail des utilisateurs augmente considérablement étant donné que leurs tâches ne sont exécutées qu'après vérification de l'environnement par l'agent-pilote. De plus, la queue de tâches centralisée permet d'appliquer de manière efficace et précise les politiques d'ordonnancement de la communauté des utilisateurs, en fonction, par exemple, des priorités de différentes activités ou de différents groupes de travail. L'organisation du travail intracommunautaire devient aussi facile que dans le cas d'un système «batch» classique d'une grappe de calcul.

Les applications doivent être adaptées pour une utilisation avec la grille de calcul. Dans beaucoup de cas, elles nécessitent des services dédiés pour mieux orchestrer le travail des jobs des utilisateurs. Le cadriciel DIRAC est spécialement conçu pour créer facilement de nouveaux services en utilisant une infrastructure unifiée de sécurité, de configuration et de surveillance. De plus, certains services développés pour des applications spécifiques ont été incorporés dans la base de logiciel de DIRAC et peuvent désormais être utilisés par d'autres communautés, par exemple catalogue des répliques distribuées développé pour la Collaboration ILC (International Linear Collider).

Le projet DIRAC fournit des interfaces utilisateurs de différents types, ce qui permet de choisir le mode de travail le plus adapté. En l'occurrence, on peut y trouver une interface riche en ligne de commandes similaires à celles de l'interface gLite, une API de programmation en langage Python et une interface Web qui permet d'accéder à quasiment toutes les fonctionnalités de DIRAC de la même manière que dans une application « desktop » graphique. Il est important de remarquer que l'interface Web de DIRAC ne fait

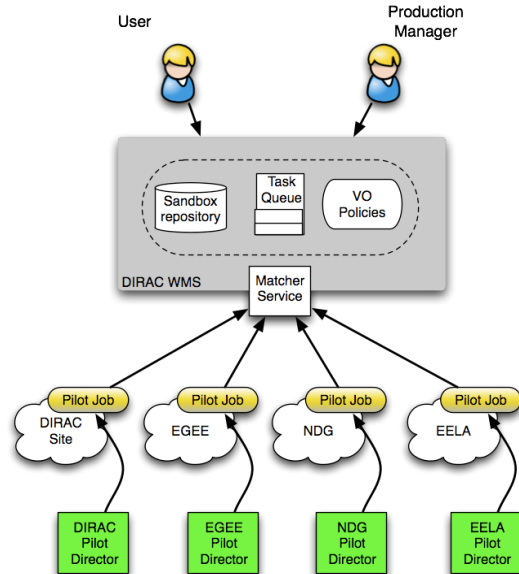


Fig.1 Système de l'ordonnancement de tâches de DIRAC avec les agents-pilotes

aucun sacrifice du point de vue de la sécurité. Toutes les opérations sont conformes aux normes de sécurité de la grille et utilisent des certificats standard pour l'authentification des utilisateurs avec des règles d'autorisation précises pour chaque opération.

Une installation des services DIRAC au Centre de Calcul (CC) de l'IN2P3 à Lyon aide les utilisateurs à accéder aux grilles de calculs. A l'origine, elle était destinée à une série de tutoriels après lesquels nombre d'utilisateurs ont commencé à l'utiliser pour leur travail pratique. Techniquement, l'installation au CC/IN2P3 est installée sur une machine virtuelle avec une configuration modeste de 30 GO d'espace disque et 8 GO de RAM. Cette configuration permet malgré tout de rendre disponible la plupart des fonctionnalités DIRAC : ordonnancement de tâches avec agents-pilotes, ainsi que leurs suivi et contrôle par les utilisateurs ; gestion de données avec un catalogue de répliques et de métadonnées; portail web pour le contrôle du fonctionnement du système et de toutes les tâches des utilisateurs. Les services de base sont aussi fournis : gestion de certificats, service de configuration et de monitoring du système, enregistrement des traces de requêtes sécurisés, etc. Tous les services DIRAC gardent leurs données dans une base MySQL. Le serveur MySQL est fourni par le CC/IN2P3 avec un support pour des sauvegardes régulières. Cette installation est suffisante pour les tutoriels et aussi pour le travail pratique des utilisateurs avec environ 1000 tâches exécutées simultanément. Grâce à la modularité du système sa capacité peut être augmentée de manière progressive. Par exemple, dans le cas de LHCb, l'installation DIRAC utilise 10 serveurs de puissance moyenne (8 cœurs, 16 GO de RAM), ce qui permet d'exécuter jusqu'à environ 50 000 tâches simultanées tout en supportant les services supplémentaires de plus haut niveau – la production et la réplication automatisées de données spécifiques à l'expérience LHCb.

D'autres installations de DIRAC ont été mises en œuvre pour des communautés particulières à CREATIS (Lyon), CPPM (Marseille), l'Université de Bordeaux, l'Université de Nice, dans des domaines variés tels que l'imagerie médicale, la sismologie, l'astrophysique, etc. Ce sont des installations de petite capacité, conçues surtout pour assister des utilisateurs locaux.

Outils, difficultés rencontrées :

DIRAC a été conçu à l'origine en tant que système pour supporter une seule communauté. Cependant, l'installation au CC de Lyon, par exemple, sert plusieurs VO avec différentes ressources et politiques d'utilisation, différents droits d'accès aux services grille, etc. C'est pourquoi le cadre de DIRAC a été étendu pour supporter différentes VO avec une seule installation. C'est pour cette raison que l'on a développé une méthode de lancement des agents-pilotes selon laquelle chaque communauté possède sa propre instance de générateur des pilotes (pilot factory). Les pilotes s'exécutent sur les ressources disponibles à leur propre VO et ne prennent que les tâches de la communauté à laquelle ils appartiennent. Cela permet de gérer les tâches provenant des utilisateurs de différentes VO tout en respectant leurs droits d'accès aux ressources et aux services de la grille.

Un autre aspect des applications autres que celles de la Physique des Hautes Energies, c'est le besoin du support d'une exécution parallèle en utilisant la technologie MPI. Ce support reste limité sur la grille. Au CC/IN2P3, la technique des agents-pilotes utilisée par DIRAC permet de lancer les applications en exécution parallèle même sur des sites qui, officiellement, ne supportent pas ce mode de travail, mais en assurent les conditions nécessaires, comme des interconnexions de nœuds de calcul suffisantes. Il y a plusieurs avantages à l'utilisation des agents-pilotes pour les tâches parallèles. Cela permet la réutilisation des groupes MPI (MPI rings) pour plusieurs travaux consécutifs, ce qui a pour effet d'augmenter l'efficacité d'exécution et diminuer la charge sur les systèmes batch des sites. De plus, cela permet d'appliquer différents types de bibliothèques MPI sans demander aux sites le support de saveurs de MPI particulières.

Résultats scientifiques :

DIRAC est un système de base pour la Collaboration LHCb mais aussi pour bon nombre d'autres collaborations dans le domaine de la Physique des Hautes Energies. De plus, les installations de DIRAC sont maintenant utilisées pour servir d'autres communautés d'utilisateurs dans le monde. En France, l'installation à CREATIS en est un bon exemple. Cette installation de DIRAC est utilisée dans le cadre du projet Virtual Imaging Platform (VIP) pour l'exécution d'applications de simulation d'images médicales sur différentes infrastructures de calcul. Dans ce projet, le système de gestion de workflows nommé MOTEUR est relié avec l'ordonnanceur des tâches DIRAC. Cette approche a permis d'exécuter les tâches non seulement sur les ressources de la grille de calcul, mais aussi sur une grappe d'ordinateurs indépendante de la grille. En 2010-2011, cette installation a été utilisée par une dizaine d'utilisateurs avec une consommation totale de plus de 10 années de CPU. Il est important de remarquer que l'utilisation de ressources de calcul hétérogène gérées par DIRAC était tout à fait transparente pour les utilisateurs de cette installation. D'autres communautés de domaines tels que l'Observation de Terre, la Physique d'Astroparticules, etc, commencent progressivement à utiliser l'infrastructure DIRAC au CC.

Perspectives :

Un bon nombre d'installations DIRAC en France ont montré l'utilité et l'efficacité du système pour une application dans différents domaines. La fonctionnalité offerte par DIRAC est appréciée par les utilisateurs car elle procure les avantages suivants :

- une efficacité d'exécution des tâches clairement plus élevée par rapport à l'infrastructure grille conventionnelle ;
- une interface d'utilisateurs plus simple et plus accessible, disponible sur les plates-formes autres que celle de la grille standard, par exemple Mac OS, avec le portail Web DIRAC ;
- possibilité d'inclure des ressources de calcul supplémentaires en provenance de grappes indépendantes ou de nuages commerciaux ;
- possibilité de développer les services pour des applications spécifiques.

Cependant, la maintenance des services DIRAC nécessite un niveau d'expertise élevé pour les administrateurs du système, il s'agit d'une tâche laborieuse et demandant beaucoup d'efforts pour la mise à jour des descriptions des ressources, pour le suivi des utilisateurs et pour la résolution de problèmes éventuels. C'est pour cette raison qu'il est prévu de réunir les efforts de différents experts pour former un groupe de support pour une seule installation puissante de DIRAC, partagée entre les utilisateurs du segment NGI français. Cela permettra d'organiser le travail des experts de la manière plus efficace et d'assurer une disponibilité des services beaucoup plus élevée que dans le cas de plusieurs installations pour des communautés relativement petites.

Références :

- [1] Projet DIRAC, <http://diracgrid.org>
- [2] A.Tsaregorodtsev et al. DIRAC3 : The New Generation of the LHCb Grid Software. Journal of Physics: Conference Series, **219**(6):062029, 2010.
- [3] A.Casajus, R.Graciani, S.Paterson, A.Tsaregorodtsev, DIRAC pilot framework and the DIRAC Workload Management System, Journal of Physics: Conference Series, **219**(6): 062049, 2010
- [4] A. Marion, G. Forestier, H. Benoit-Cattin, S. Camarasu-Pop, P. Clarysse, R. Ferreira da Silva, B. Gibaud, T. Glatard, P. Hugonnard, C. Lartzien, H. Liebgott, J. Tabary, S. Valette, and D. Friboulet. Multimodality medical image simulation of biological models with the Virtual Imaging Platform (VIP). In IEEE CBMS 2011, Bristol, UK, 2011.